

AD-A250 401



## INFORMATION PAGE

Form Approved  
OMB No. 0704-0188

ited to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering the collection of information. Send comments regarding this burden estimate or any other aspect of this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson  
Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

RT DATE

## 3. REPORT TYPE AND DATES COVERED

FINAL 01 Jun 88 TO 31 Jan 92

## 4. TITLE AND SUBTITLE

HUMAN IMAGE UNDERSTANDING

DTIC

SECRET  
MAY 19 1992

## 5. FUNDING NUMBERS

AFOSR-88-0231

PE 61102F

PR 2313

TA A5

## 6. AUTHOR(S)

Dr Irving Biederman

## 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Psychology  
University of Minnesota  
Elliott Hall  
75 East River Road  
Minneapolis, MN 554558. PERFORMING ORGANIZATION  
REPORT NUMBER

## 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

DR JOHN F. TANGNEY  
AFOSR/NL  
Building 410  
Bolling AFB DC 20332-644810. SPONSORING/MONITORING  
AGENCY REPORT NUMBER

92-12940



## 11. SUPPLEMENTARY NOTES

## 12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

## 12b. DISTRIBUTION CODE

## 13. ABSTRACT (Maximum 200 words)

This report summarizes the major research accomplishments performed under AFOSR Grant 88-0231, HUMAN IMAGE UNDERSTANDING. An extensive series of experiments assessing the visual priming of briefly presented images indicate that the visual representation that mediates real-time object recognition specifies neither the image edges or vertices nor an overall model of the object but an arrangement of simple volumes (or geons) corresponding to the object's parts. This representation can be activated with no loss in efficiency when the image is projected onto the retina at another position, size, or orientation in depth from when originally viewed. Consideration of these invariances suggests a computational basis for the evolution of two extrastriate visual systems, one for recognition and the other subserving motor interaction. It may be possible to assess the functioning of these systems behaviorally, that is, to split the cortex horizontally, through a comparison of performance on naming and episodic memory tasks. We have developed a neural network model (Hummel & Biederman, 1992) that captures the essential characteristics of human object recognition performance. The model takes a line drawing of an object as input and generates a structural description which is then used for object classification. The model's

## 14. SUBJECT TERMS

## 15. NUMBER OF PAGES

## 16. PRICE CODE

17. SECURITY CLASSIFICATION  
OF REPORT

(U)

18. SECURITY CLASSIFICATION  
OF THIS PAGE

(U)

19. SECURITY CLASSIFICATION  
OF ABSTRACT

(U)

## 20. LIMITATION OF ABSTRACT

(U)

capacity for structural description derives from its solution to the dynamic binding problem of neural networks: Independent units representing an object's parts (in terms of their shape attributes and interrelations) are bound temporarily when those attributes occur in conjunction in the systems input. Temporary conjunctions of attributes are represented by synchronized activity among the units representing those attributes. Specifically, the model induces temporal correlation in the firing of activated units to: a) parse images into their constituent parts; b) bind together the attributes of a part; and c) determine the relations among the parts and bind them to the parts to which they apply. Because it conjoins independent units temporarily, dynamic binding allows tremendous economy of representation, and permits the representation to reflect an object's attribute structure. The model's recognition performance conforms well to recent results from shape priming experiments. Moreover, the manner in which the model's performance degrades due to accidental synchrony produced by an excess of phase sets suggests a basis for a theory of visual attention.

## FINAL PROGRESS REPORT

Submitted to: Air Force Office of Scientific Research  
AFOSR/NL  
Building 410  
Bolling A.F.B., D.C. 20332-6448  
  
Dr. John Tangney, Program Manager

Title: HUMAN IMAGE UNDERSTANDING

Grant No. AFOSR-88-0231

Date April 17, 1992

Submitting  
Organization: University of Minnesota  
Office of Research Administration  
1919 University Ave.  
St. Paul, MN 55104

PI: Professor Irving Biederman  
(213) 740-6094

Irving Biederman  
134-28-9789

Department of Psychology  
University of Minnesota  
Elliott Hall  
75 East River Road  
Minneapolis, MN 55455

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

**Final Progress Report for AFOSR Grant No. AFOSR-88-0221**  
**Psychophysical Analyses of Perceptual Representations**

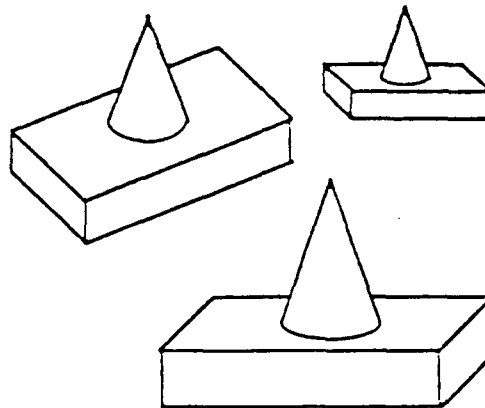
**Abstract**

This report summarizes the major research accomplishments performed under AFOSR Grant 88-0231, HUMAN IMAGE UNDERSTANDING. An extensive series of experiments assessing the visual priming of briefly presented images indicate that the visual representation that mediates real-time object recognition specifies neither the image edges or vertices nor an overall model of the object but an arrangement of simple volumes (or geons) corresponding to the object's parts. This representation can be activated with no loss in efficiency when the image is projected onto the retina at another position, size, or orientation in depth from when originally viewed. Consideration of these invariances suggests a computational basis for the evolution of two extrastriate visual systems, one for recognition and the other subserving motor interaction. It may be possible to assess the functioning of these systems behaviorally, that is, to split the cortex horizontally, through a comparison of performance on naming and episodic memory tasks. We have developed a neural network model (Hummel & Biederman, 1992) that captures the essential characteristics of human object recognition performance. The model takes a line drawing of an object as input and generates a structural description which is then used for object classification. The model's capacity for structural description derives from its solution to the *dynamic binding problem* of neural networks: Independent units representing an object's parts (in terms of their shape attributes and interrelations) are bound *temporarily* when those attributes occur in conjunction in the systems input. Temporary conjunctions of attributes are represented by synchronized activity among the units representing those attributes. Specifically, the model induces temporal correlation in the firing of activated units to: a) parse images into their constituent parts; b) bind together the attributes of a part; and c) determine the relations among the parts and bind them to the parts to which they apply. Because it conjoins independent units temporarily, dynamic binding allows tremendous economy of representation, and permits the representation to reflect an object's attribute structure. The model's recognition performance conforms well to recent results from shape priming experiments. Moreover, the manner in which the model's performance degrades due to accidental synchrony produced by an excess of phase sets suggests a basis for a theory of visual attention.

This document summarizes the major research contributions to derive from Air Force Office of Scientific Research Grant 88-0231, *Human Image Understanding*, to Irving Biederman. The initial section presents an overall summary, followed by more detailed descriptions, largely taken from the abstracts of the published reports from this work. These published reports have been submitted to the monitor under separate cover. The final section lists the publications, presentations, and recognition that the research supported by the grant has received.

### SUMMARY OF RESEARCH CONTRIBUTIONS

Consider figure 1. We can appreciate that the three images represent the same (unfamiliar) object, despite substantial differences in size, position, and orientation in depth. We will refer to these variations as variations in viewpoint.



*Fig. 1.* The above shape is readily detectable as constant across the three views despite its being unfamiliar.

The subjective equivalence of the three images in figure 1 is not illusory. Recent object priming studies on this project have established that, indeed, the speed of object recognition, as assessed by visual priming of naming latencies, is invariant with translation, scale, and orientation in depth (up to parts occlusion) (Biederman & Cooper, in press, a, b [Appendices A & B]; Biederman, 1991; Gerhardstein & Biederman, 1991). A weak form of invariance that would imply that human observers could appreciate that the three objects depicted in figure 1 are the same shape. Casual viewing, of the kind invited in the first paragraph of this section, is sufficient to document that such invariance can be achieved.

There is a strong form of invariance, however, concerning the time required to achieve the equivalence, that surprisingly, has only rarely been tested in visual shape recognition. That is, the considerable facilitation in the naming RTs (reaction times) and error rates in the naming of brief, masked pictures of objects on a second block of trials, presented several minutes after they were named on a first block, is unaffected by a change in the position of the object relative to fixation (either left-right or up-down), its size, or its orientation in depth. That a considerable portion of the priming is visual (and not just a function of activation of the name or entry-level concept) is evidenced by a large reduction in priming for a same name, different shaped exemplar, as when a grand piano is shown on the second block when initially an upright piano was viewed.

These invariances are so fundamental to object recognition that theory in this domain consists largely of explaining how they could come about. On computational grounds, the invariances seem entirely reasonable in that the alternative, a separate representation of an object for each of its image manifestations, would require a prohibitively large number of representations. The invariance in recognition *speed*, moreover, is inconsistent with the hypothesis (such as that advanced by Ullman, 1989) that recognition is achieved through template transformations for translating, scaling, or rotating an image or template so as to place the two in correspondence, as such transformations would (presumably) require time for their execution, not to mention the formidable initial problem of selecting the appropriate transformation to apply to an unknown image.

Now consider figure 2, in which we are to judge which one of the three stimuli is not like the others. We readily select 3 as different, yet there is a much greater difference in contour (as assessed by the number of mismatching pixels in the best match) between the image of object 2 and the other two images in that 2's brick is more elongated than the bricks of the other two objects (whose bricks are identical). Objects 2 and 3 have identical cones, so the difference is in the aspect ratio of their bricks. This demonstration suggests that relatively small differences in contour that produce a qualitative difference--whether the tip of the cone is pointed or rounded in the example--can have a more noticeable effect on classification than larger differences in a metric property, such as aspect ratio, which varies with viewpoint. Our interpretation of qualitative is "viewpoint invariant" and the empirical work described in this report is, to a large extent, concerned with exploring how viewpoint invariance in visual shape recognition performance can be achieved.<sup>1</sup>

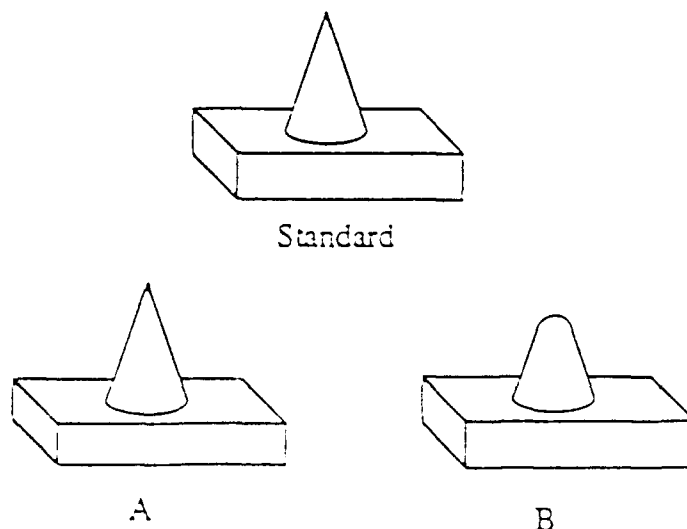


Fig 2. Object A is judged to be more similar to the standard than B, but B is a closer match for a template model.

We have developed a theory to account for this capacity, *Recognition-by-Components* (RBC) which posits that, for purposes of entry level recognition, objects are represented as an arrangements of convex volumetric primitives such as bricks, wedges, cylinders, cones, lemons,

<sup>1</sup>"Viewpoint invariance" can refer to: a) stability of certain kinds of contour information with changes of viewpoint, and b) the lack of an effect on performance from changes in viewpoint. The context should disambiguate which sense is intended.

and their singly concave curved-axes counterparts, such as a cylinder with a curved axis (Biederman, 1987), as illustrated in Figure 3. There are 24 primitives, called *geons*, in the current version of the theory and they have the property that they can be distinguished from a general viewpoint. For example, from almost any viewpoint one would be able to distinguish a brick from a cylinder. Once two or three geons and their relations are specified, then almost any image of an object can be recognized as an instance of its entry level class.

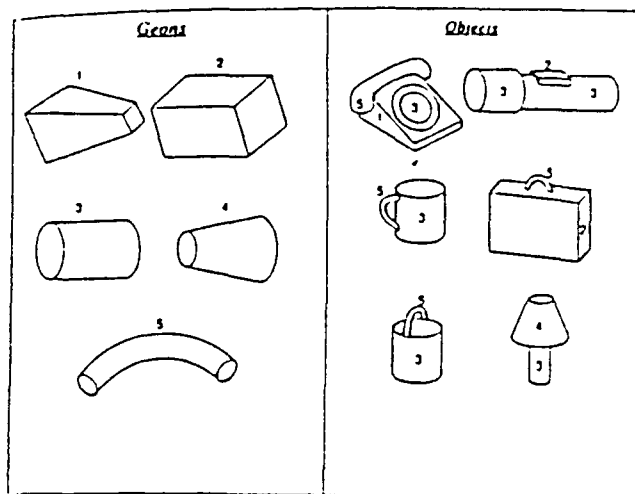


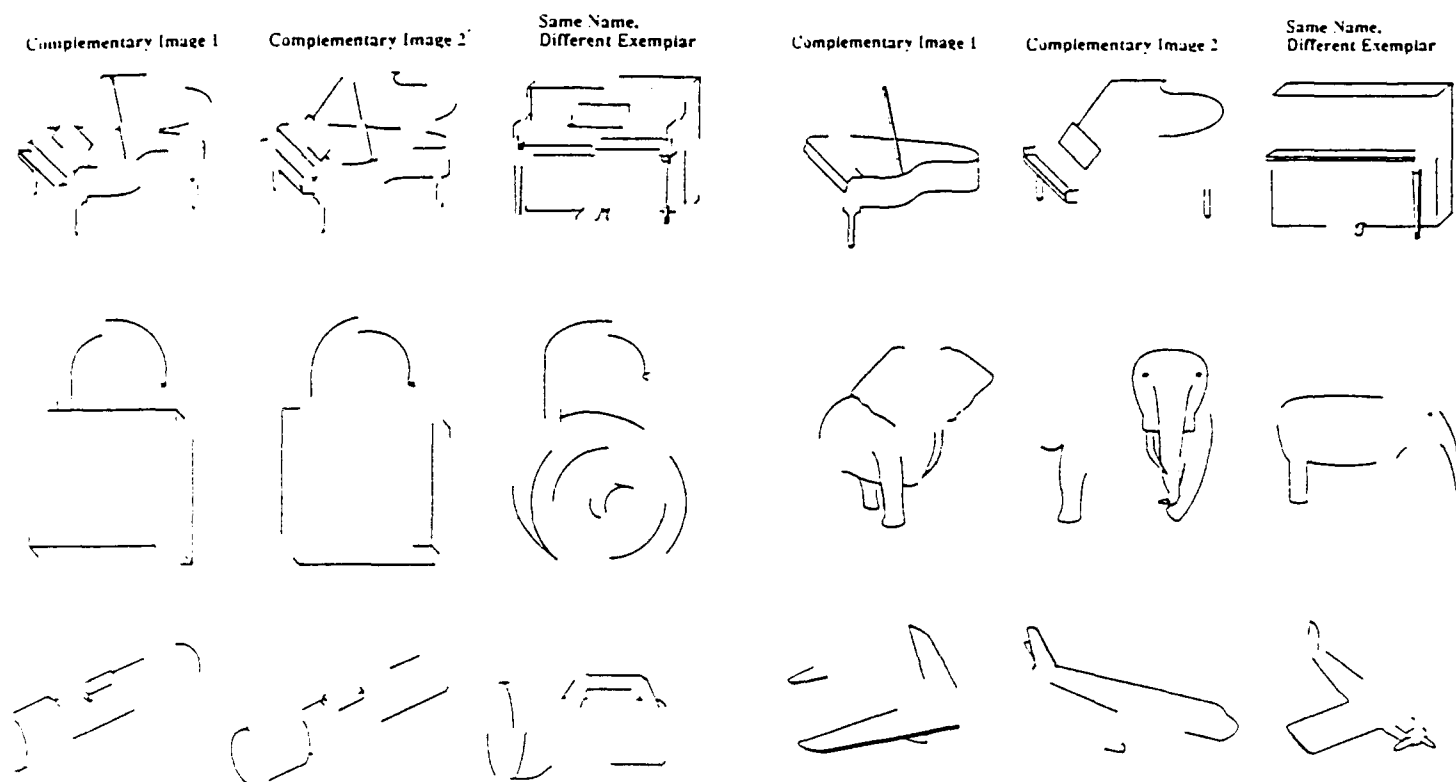
Fig. 3 . A given view of an object can be represented as an arrangement of simple primitive volumes, or geons, of which five are shown here. Only two or three of the geons are needed to specify an object.

### *Parts-based recognition*

What evidence is there that that object recognition is (simple) parts-based? Consideration of this question requires explication of alternatives to parts-based representation. Two have been proposed, templates and lower level features. Some of the evidence against templates are the robustness of recognition when an object is presented at a novel orientation in depth, or with some of its parts removed, or with the addition of irrelevant parts.

Biederman and Cooper (1991) recently reported a more direct test of the alternatives. They used picture priming tasks to assess whether the facilitation of naming RTs and accuracy on a second block of object pictures is a function of the repetition of the object's: a) image features (viz., vertices and edges), b) the object model (e.g., that it is a grand piano), or c) a representation intermediate between a) and b) consisting of convex or singly concave components of the object, roughly corresponding to the object's parts. Subjects viewed pictures with half their contour removed by either a) deleting every other image feature from each part (as shown in Fig. 4), or b) half the components (as shown in Fig. 5). On a second (primed) block of trials, subjects saw: a) the identical image that they viewed on the first block, b) the complement which had the missing contours, or c) a same name, different exemplar of the object class (e.g., a grand piano when an upright piano had been shown on the first block). With deletion of features, speed and accuracy of naming identical and complementary images were equivalent, indicating that none of the priming could be attributed to the features actually present in the image. Performance with both types of image enjoyed an advantage over the different exemplars, establishing that the priming was visual, rather than verbal or conceptual. With deletion of the components, performance with identical images was much better than their complements. The latter were equivalent to the different exemplars, indicating that *all* the visual priming of an image of an object can be modeled in terms

of a representation of its components (in specified relations). Alternative explanations are still somewhat viable and a portion of the proposed effort is directed toward their assessment.



*Figures 4 (left) showing feature deletion and 5 (right) showing parts deletion.* First two columns in each panel: Complementary pairs of images created by deleting every other edge and vertex from each geon (Fig. 4) or half the parts (Fig. 5). Each member of a complementary pair had half the contour so that if the members of a pair were superimposed, the composite would make for an intact picture without any overlap in contour. Assuming that the image in the left column was shown on the first block, the same image on the second block would be an instance of identical priming, the middle image would be complementary priming, and the right would be a different exemplar (same name) control. For images of the type shown in Fig. 4., identical and complementary conditions produced equivalent priming, both of which were greater (in priming) than the different exemplar condition. For images of the type in Fig. 5., more priming was associated with the identical images than either the complementary or different exemplar images, which did not differ from each other.

#### *Neural Net Implementation of RBC*

These presumed characteristics of human shape recognition--invariant, parts-based representations--have provided the goals for a neural net implementation of RBC that takes as its input a line drawing of an object's orientation and depth discontinuities and as output activates a unit representing a structural description that is invariant with position, size, and orientation in



depth (Hummel & Biederman, in press [Appendix C]). Figure 6 shows the overall architecture of the model (called, JIM).

JIM's capacity for structural description derives from its solution to the *dynamic binding problem* of neural networks: Independent units representing an object's parts (in terms of their shape attributes and interrelations) are bound *temporarily* when those attributes occur in conjunction in the systems input. Temporary conjunctions of attributes are represented by synchronized (or phase locked) oscillatory activity among the units representing those attributes. Specifically, the model uses phase locking to: a) parse images into their constituent parts; b) bind together the attributes of a part; and c) determine the relations among the parts and bind them to the parts to which they apply. Because it conjoins independent units temporarily, dynamic binding allows tremendous economy of representation, and permits the representation to reflect the attribute structure of the shapes represented.

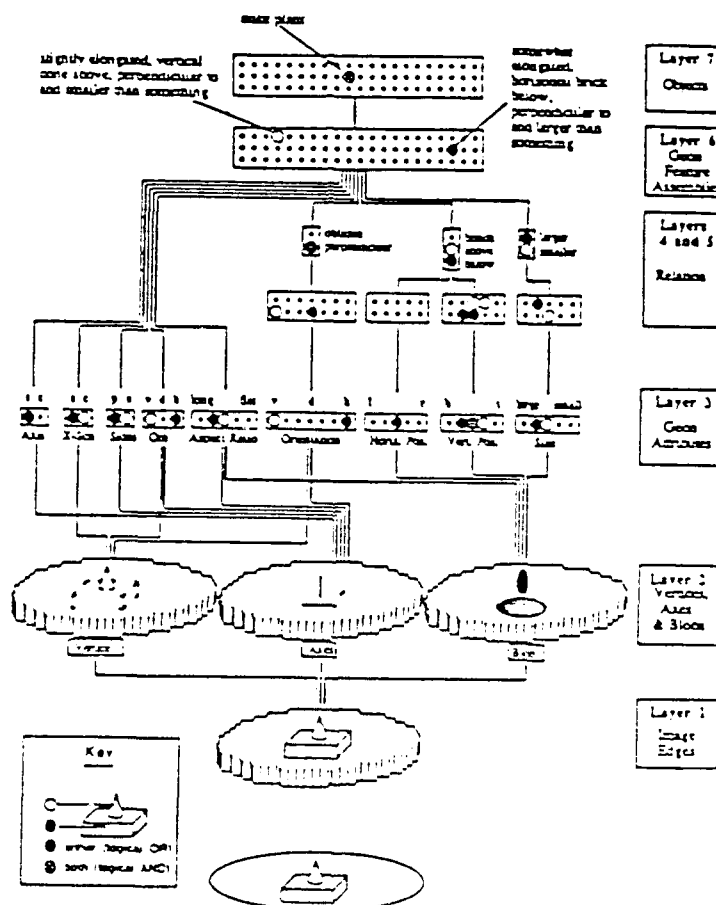
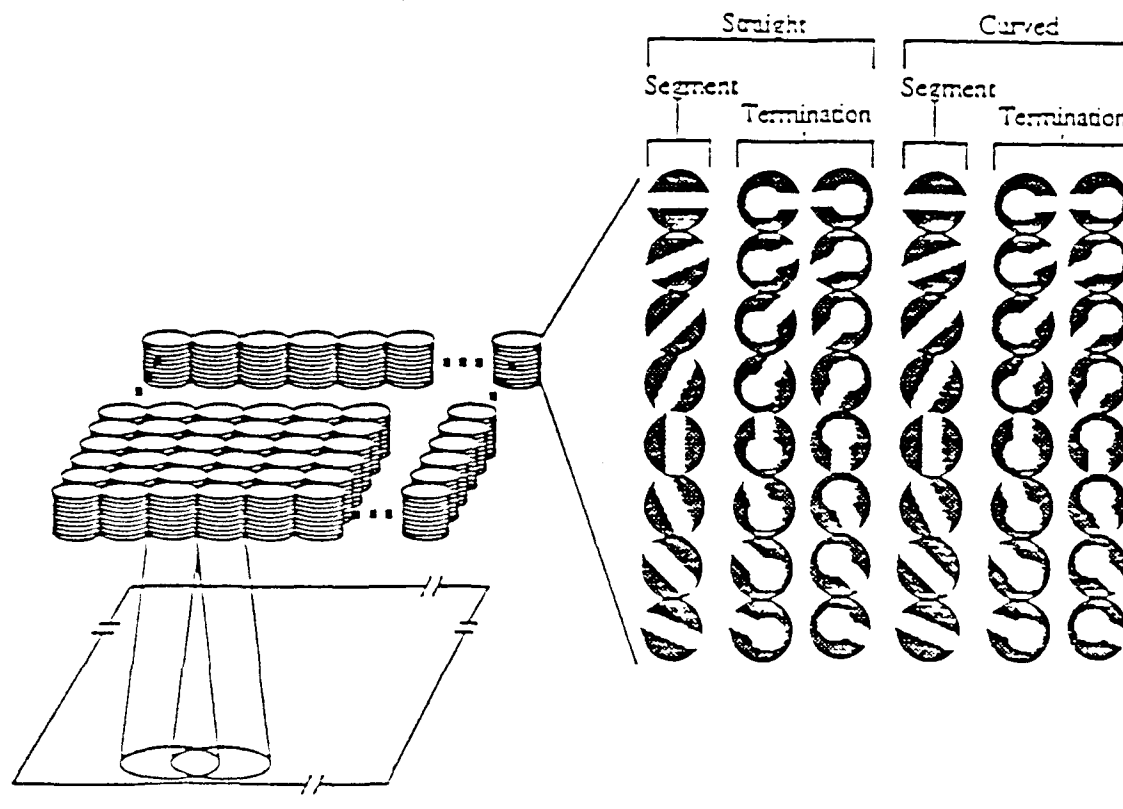


Fig. 6. An overview of the net's architecture indicating the representation activated at each layer. In L (layer) 3 and above, large circles indicate cells activated in response to the image and dots indicate inactive cells. Cells in L1 represent the edges (specifying discontinuities in surface orientation and depth). L2 represents the vertices, axes, and blobs defined by conjunctions of edges in L1. L3 represents the geons in terms of their defining attributes (Axis, straight or curved), Cross section (straight or curved) and sides (parallel or not parallel) as well as coarse coding of metric attributes of the geons. L4 and L5 represent the relative relations among geons. Cells in L6 respond to conjunctions of cells in L3 and L5, and cells in L7 respond to conjunctions of L6 cells.

*Brief Overview of JIM.* As shown in Figure 7, layer 1 (L1) is a highly simplified version of V1 in which a 22 X 22 array of spatially arranged columns (roughly analogous to V1 hypercolumns), each with 48 cells that respond to local lines (differentially to straight vs. curved and end-stopped vs. segments that extend through the receptive field) at various orientations. A second layer contains units that respond to vertices at various orientations (activated by the end stopped units in L1), axis of surfaces, and the general mass (blob) of a volume. Binding is initiated at these first two layers by Fast Enabling Links (FELs), connections between pairs of cells that result in phase locking the outputs of activated cells that are collinear (or cocircular), closely parallel or coterminating. For example, the various collinear segment cells activated by a line with a length greater than the receptive field diameter of those cells will all fire in synchrony.

The model's first layer is  
divided into 22 X 22 locations.

At each location there are  
48 cells.



*Figure 7.* Detail of the model's first layer. Image edges are represented in terms of their location in the visual field, orientation, curvature, and whether they terminate within the cell's receptive field or pass through it.

The units in L2 activate 55 units in L3 that provide an invariant representation of the object's geons and the characteristics of these geons (viz., aspect ratio and absolute orientation [vertical, horizontal, or oblique]). The phase locking in the third layer is maintained so that all the cells that represent a particular geon, say the brick in Figure 7, will fire synchronously but out of phase with the firing of another geon in that object, say the cone in Figure 7. Outputs from the

size, location and orientation units in L3 activate units in L4 and L5 that represent invariant relations between pairs of geons, such as relative position (above, below, side of), relative size (larger than, smaller than, equal to), or relative orientation (perpendicular, parallel, or diagonal to). The phase locking is maintained through these layers as well so that simultaneously arriving L3 and L5 outputs will recruit a given *geon feature assembly* (GFA) cell in L6. Such cells will represent a given geon, its attributes, and its pairwise relations to other geons, e.g., that a brick (actually a part with a straight cross section, straight axis, and constant sized cross section), that is horizontal, wider than it is tall, below, larger than, and perpendicular to something else is present in the object. Closely firing L6 cells will recruit a given L7 cell which will represent a given object.

*Locus of priming.* In the context of the model, where is the locus of visual object priming? The absence of object model priming (as evidenced by the absence of priming between complements with different parts) suggests that it cannot be attributed to residual activation of L7 cells. The failure to find any contributions from reinstatement of edges and vertices argues against the substrate existing at L1 or the vertex units in L2. Moreover, because the same units in L1 to L5 are used repeatedly for different objects, activation of any one unit of a particular object would be readily overwritten by the activation of that and other units in that layer by other objects. In general, the first five layers would presumably be set beyond (if not before) infancy, so it would be unlikely for priming to have a noticeable effect at these early stages. We (Cooper & Biederman, 1991) tested this proposition directly by presenting, immediately prior to the presentation of an object, the single largest geon from that object. Such a prime would contain the specifications of a geon of the object, its absolute orientation and aspect ratio but none of the relations, such as TOP-OF, LARGER-THAN, and PERPENDICULAR TO. Compared to control trials in which the prime was not contained in the object or no prime was presented, no priming was observed. In fact, before this experiment was done, it was apparent that this result was predicted from JIM. The reason for this is that a geon feature assembly cell in L6 has to have a high "vigilance parameter" (or sharp tuning function) if it is to distinguish among objects that contain the same geons. In particular, without the inputs from the relation units, L6 units would be activated by similar geons from competing objects. An analogy can be made through a gedanken experiment in which one might attempt to prime five letter words with a single letter. No priming would be expected; not because letters are irrelevant to words, but because distinctiveness requires specification of both a particular letter (or spelling pattern which consists of a particular group of letters) at a given *position* in a letter sequence.

Priming would thus be localized at three possible sites: a) the weight matrix for L3 & L5 --> L6 would be the earliest locus where priming should be manifested, b) activation of L6 units, and/or c) the L6 --> L7 weight matrix.

## SUMMARY OF INDIVIDUAL PROJECTS

The summary of the individual research projects is divided into three major sections. The research described in Section I employed priming to study the form of the representation (Part A) and the invariances (Part B). Several methodological studies have also been performed on the technique itself. Section II describes a major effort has centered on the development of a neural net implementation of RBC. Section III describes research designed to explore the cortical implementation of object recognition, including some new work on patient populations. Major references are provided after each abstract.

## **I. Studies of Priming:**

### **A. Nature of the Representation**

#### **1. Priming Contour-Deleted Images: Evidence for Intermediate Representations Mediating Visual Object Recognition Rather than Specific Contours (edges and vertices) or Subordinate Models.**

The speed and accuracy of perceptual recognition of a briefly presented picture of an object is facilitated by its prior presentation. Picture priming tasks were used to assess whether the facilitation is a function of the repetition of : (a) the object's image features (viz., vertices and edges), (b) the object model (e.g., that it is a grand piano), or (c) a representation intermediate between (a) and (b) consisting of convex or singly concave components of the object, roughly corresponding to the object's parts. Subjects viewed pictures with half their contour removed by deleting either: (a) every other image feature from each part, or (b) half the components. On a second (primed) block of trials, subjects saw: (a) the identical image that they viewed on the first block, (b) the complement which had the missing contours, or (c) a same name-different exemplar of the object class (e.g., a grand piano when an upright piano had been shown on the first block). With deletion of features, speed and accuracy of naming identical and complementary images were equivalent, indicating that none of the priming could be attributed to the features actually present in the image. Performance with both types of image enjoyed an advantage over that with the different exemplars, establishing that the priming was visual, rather than verbal or conceptual. With deletion of the components, performance with identical images was much better than that with their complements. The latter were equivalent to the different exemplars, indicating that all the visual priming of an image of an object is through the activation of a representation of its components in specified relations. In terms of a recent neural net implementation of object recognition (Hummel & Biederman, In press), the results suggest that the locus of object priming may be at changes in the weight matrix for a geon assembly layer, where units have self-organized to represent combinations of convex or singly concave components (or geons) and their attributes (e.g., aspect ratio, orientation and relations with other geons such as TOP-OF). The results of these experiments provide evidence for the psychological reality of intermediate representations in real-time visual object recognition.

#### *Reference*

Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, 23, 393-419.

#### **2. Pattern Goodness Can be Understood as the Working of the Same Processes that Produce Invariant Parts for Purposes of Object Recognition.**

Pattern goodness, or *pragnanz*, has been a subject of study and theorizing for over half a century but its role in vision remains uncertain. The traditional theoretical dispute as to whether goodness reflects a tendency for perception to derive the simplest interpretation of the stimulus versus the most frequently occurring pattern in the environment is probably unresolvable in the absence of a theory that defined what a stimulus was (particularly one projected from a three dimensional object), so that its likelihood could be determined, and the manner in which constraints toward simplicity could or could not be regarded as something extractable from the regularities of images. We argue that it is likely that goodness effects are epiphenomenal, reflecting the operation of perceptual mechanisms designed to infer a three dimensional world from parts segmented from a two dimensional image and provide descriptions of objects that can be recognized from a novel viewpoint or that are partially occluded. These perceptual mechanisms are scale sensitive and include processes for viewpoint-invariant edge characterization, segmentation, and the activation of shape representations.

### *Reference*

Biederman, I., Hilton, H. J., & Hummel, J. E. (1991). Pattern goodness and pattern recognition, Ch. 5, Pp. 73-95. In J. R. Pomerantz & G. R. Lockhead (Eds.) *The Perception of Structure*. Washington, D.C.: APA.

### **3. Single Volumes are Insufficient Primes--Relations are Needed as Well.**

Subjects are faster to name an object picture with a basic level name which they have previously named. Biederman and Cooper (1991) have shown that the perceptual portion of this *priming effect* does depend on the repetition of the image features (edges and vertices) present in the original image or of the overall object model, but rather involves simple components, often corresponding to an object's parts, intermediate between these two representations. Two experiments were conducted to determine the representational level at which priming occurs. Subjects named objects that could be preceded by a single volume prime (which could either be present or absent in the object) or a neutral line. No effect of prime type was found on object naming RTs or errors even when the objects' identities were made salient by displaying them beforehand. The results support a representational level specifying an object's convex components and their relations as the locus of priming.

### *Reference*

Cooper, E. E., & Biederman, I. (1991). Priming objects with single volumes. Submitted for publication.

## **B. INVARIANCE**

### **1. Size Invariance in Visual Object Priming**

*Abstract.* The magnitude of priming resulting from the perception of a briefly presented picture of an object in an earlier trial block, as assessed by naming reaction times (RTs), was found to be independent of whether the primed object was presented at the same or a different size as when originally viewed. In contrast, RTs and error rates for "same" responses for old-new shape judgments were very much increased by a change in object size from initial presentation. We conjecture that this dissociation between the effects of size consistency on naming and old-new shape recognition may reflect the differential functioning of two independent systems subserving object memory: one for representing the shape of an object and the other for representing its size, position, and orientation (*metric* attributes). With allowance for response selection, object naming RTs may provide a relatively pure measure of the functioning of the shape system. Both the shape and metric systems may affect the feelings of familiarity that govern old-new episodic shape judgments. A comparison of speeded naming and episodic recognition judgments may provide a behavioral, noninvasive technique for determining the neural loci of these two systems.

### *Reference*

Biederman, I., & Cooper, E. E. (1992). Scale invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, In press.

## 2. Translational and Reflectional Invariance in Visual Object Priming

The magnitude of priming on naming reaction times and on the error rates, resulting from the perception of a briefly presented picture of an object approximately 7 min before the primed object, was found to be independent of whether the primed object was originally viewed in the same hemifield, left-right or upper-lower, or in the same left-right orientation. Performance for same-name, different-exemplar images was worse than for identical images, indicating that not only was there priming from block one to block two, but that some of the priming was visual, rather than purely verbal or conceptual. These results provide evidence for complete translational and reflectional invariance in the representation of objects for purposes of visual recognition. Explicit recognition memory for position and orientation was above chance, suggesting that the representation of objects for recognition is independent of the representations of the location and left-right orientation of objects in space.

### Reference

Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, In press.

## 3. 3D Orientation Invariance

Several recent reports have documented extraordinary difficulty in the recognition of images of certain kinds of unfamiliar 3D objects from a novel orientation in depth (Edelman & Bulthoff, 1991; Rock & DiVita, 1987; Tarr, 1989). The difficulty at specific orientations can be greatly reduced with practice at those orientations. If generally true, such a result would support the contention that the capacity to recognize everyday objects is a consequence of familiarity over a variety of viewpoints, in which separate visual representations (templates) are created for each experienced viewpoint. Such a theory would stand in contrast to invariant-parts theories of basic level object recognition such as RBC (Biederman, 1987; Hummel & Biederman, 1992) which assume that a viewpoint invariant structural description (up to parts occlusion and accretion) can be created from a single view of many objects, whatever their familiarity. Three experiments are reported. The first revealed complete viewpoint invariance in the visual priming of novel images of familiar objects in that changes of up to 135° in depth resulted in virtually no reduction in the magnitude of facilitation of naming RTs, as shown in Figure 8. That it was *visual* priming and not just name or concept priming was evidenced by the advantage of identical images over same name, different shaped exemplars. The second experiment showed that name priming could be reduced if there was a change in the part descriptions (so that some parts were deleted and other parts accreted) from priming to primed trials. The third experiment employed unfamiliar objects composed of novel arrangements of volumes, shown in Figure 9. Same-different judgments of sequentially presented images showed little cost of rotation in depth as long as the same invariant parts description could be activated (Figure 9). Together these results suggest that depth invariance can be readily achieved if the different stimuli activate *distinctively different* and *viewpoint invariant* (e.g., geon) representations. These two specifications may constitute the defining perceptual conditions for the formation of basic (or entry) level categories.

### References

- Biederman, I., & Gerhardstein, P. C. (1992). Recognizing depth-rotated objects: Evidence for 3D viewpoint invariance. Submitted to the *Journal of Experimental Psychology: Human Perception and Performance*.
- Gerhardstein, P. C., & Biederman, I. 3D orientation invariance in visual object recognition. Paper presented at the Meetings of ARVO, Sarasota, FL., May 1991.

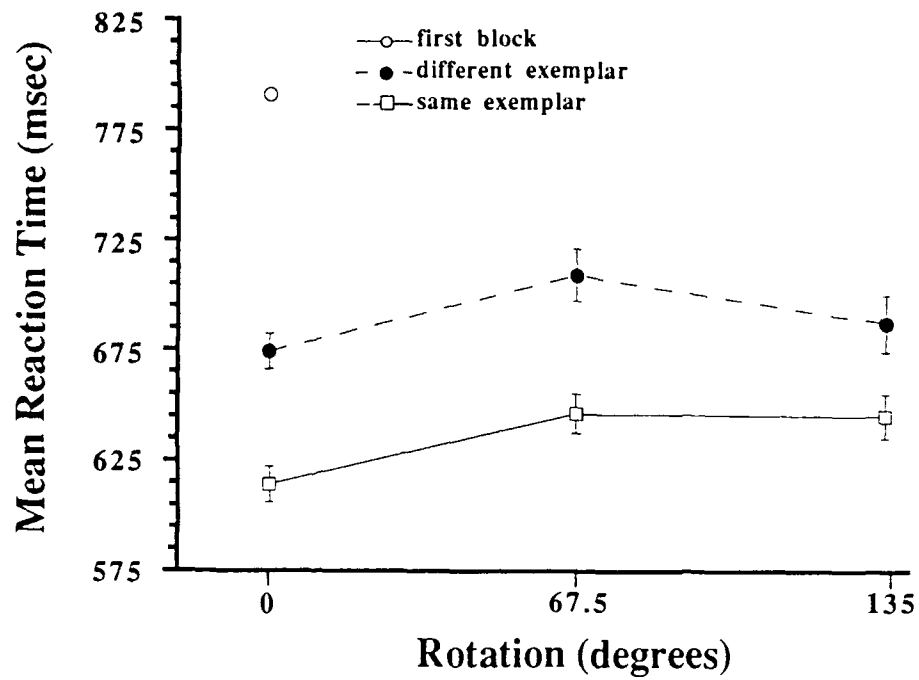


Figure 8. Mean correct RTs as a function of orientation change in Experiment I. Error bars denote SEs.

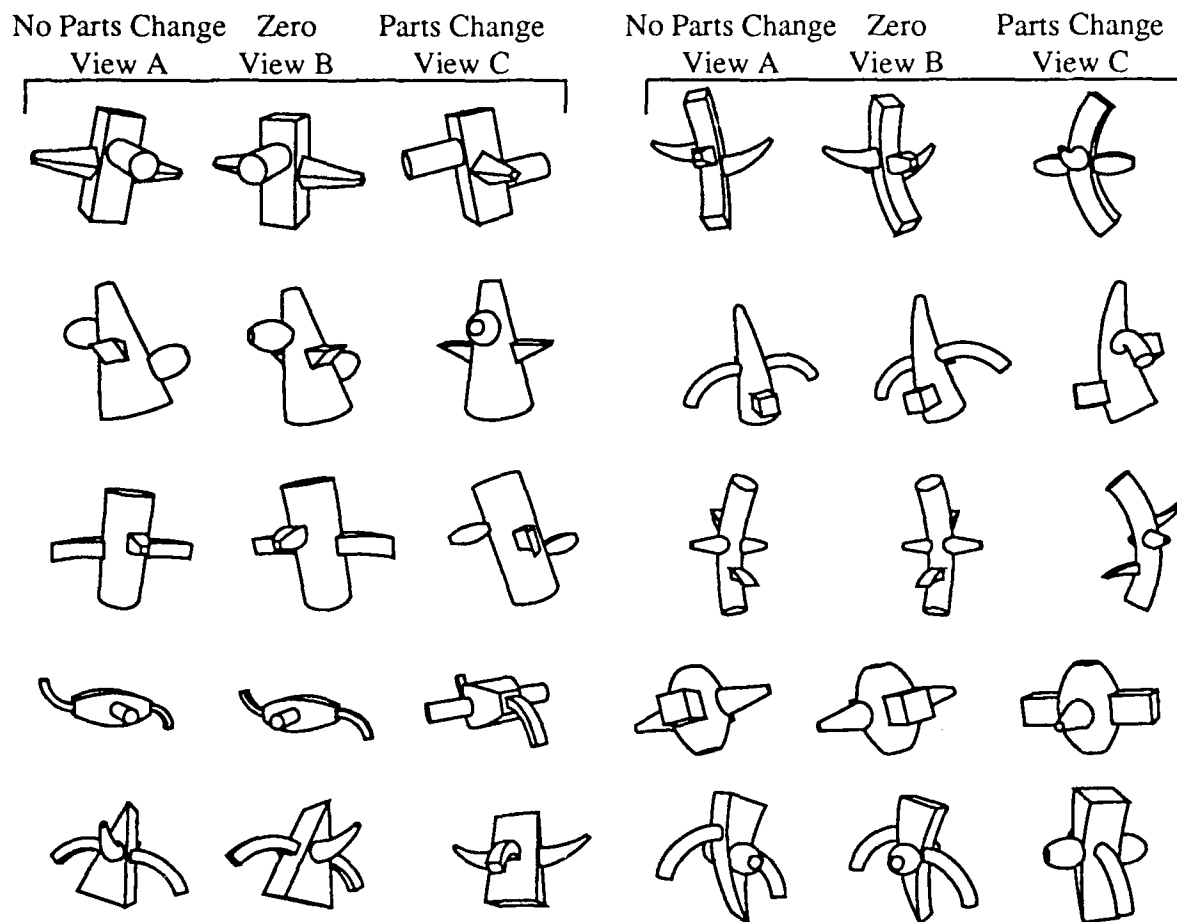


Figure 9. The ten objects used in the Experiment III of Biederman and Gerhardstein (1992). The No Parts Change and Parts Change views are all rotations of  $45^\circ$  in depth from the Zero view. Note that no object contains a part unique to that object, and relations between object parts are the same for all objects.



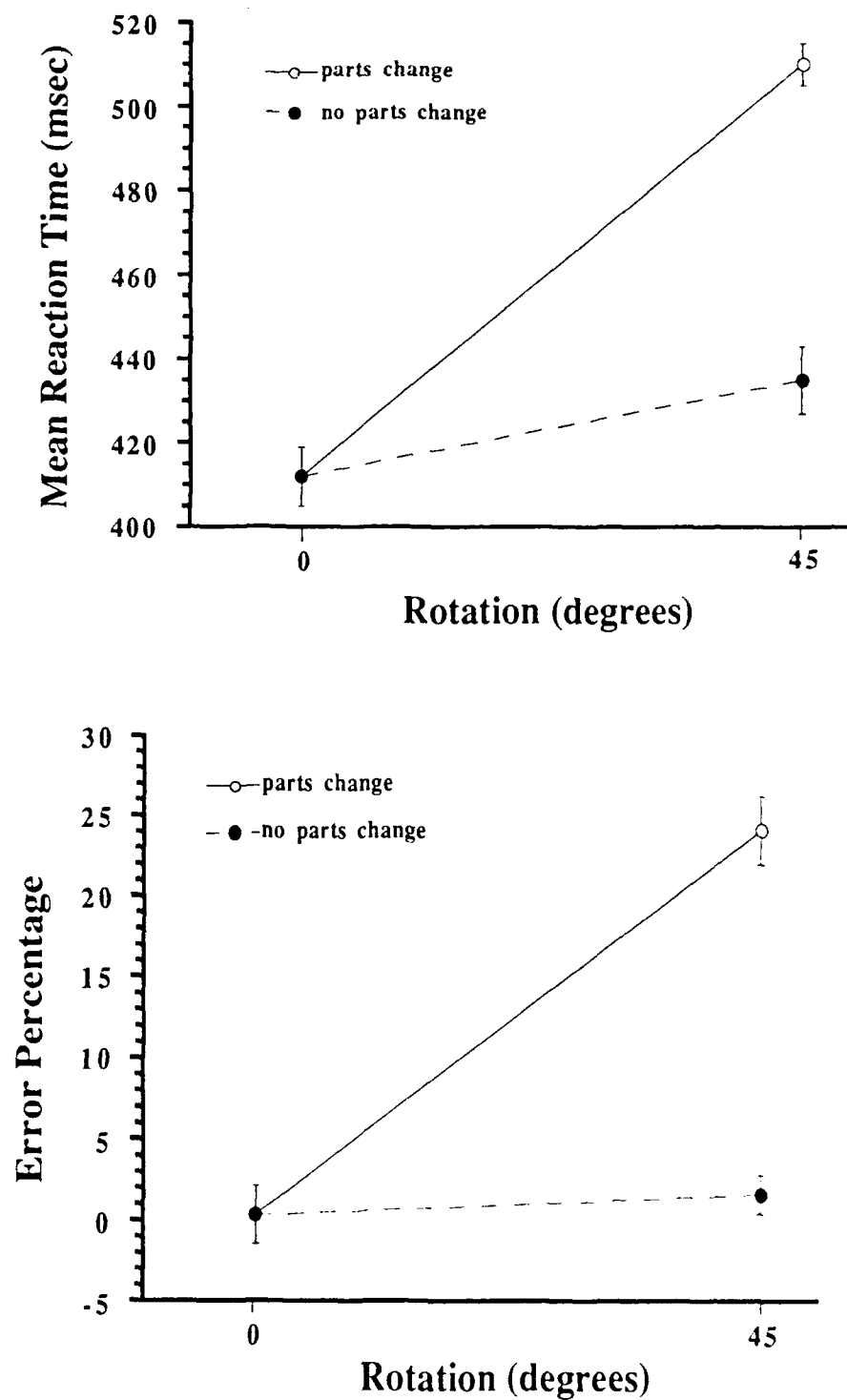


Figure 10. Mean correct RTs and error rates as a function of the amount of angular change between the first and second exposure on a trial in Experiment III. Error bars denote SEs.

#### 4. SHAPE INVARIANCE: A REVIEW AND FURTHER EVIDENCE

**Abstract.** Phenomenologically, human shape recognition appears to be invariant with changes of orientation in depth (up to parts occlusion), position in the visual field, and size. It is possible that these invariances are achieved through the application of transformations such as rotation, translation, and scaling of the image so that it can be matched metrically to a stored template. Presumably, such transformations would require time for their execution. We describe recent priming experiments in which the effects of a prior brief presentation of an image on its subsequent recognition is assessed. The results of these experiments indicate that the invariance is complete: The magnitude of visual priming (as distinct from name or basic level concept priming) is not affected by a change in position, size, orientation in depth, or lines and vertices, as long as representations of the same components can be activated. An implemented seven layer neural network model (Hummel & Biederman, in press) that captures these fundamental properties of human object recognition is described. Given a line drawing of an object, the model activates a viewpoint-invariant structural description of the object specifying its parts and their interrelations. Visual priming is interpreted as a change in the connection weights for the activation of: a) cells representing *geon feature assemblies* (GFAs), cells that conjoin the output of units that represent invariant, independent properties of a single geon and its relations (such as its type, aspect ratio, relations to other geons), or b) a change in the connection weights by which several GFAs activate a cell representing an object.

##### *Reference*

Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: A review and additional evidence. *Canadian Journal of Psychology*, In press.

#### C. PRIMING METHODOLOGY

##### 1. Name and concept priming

Researchers using object naming latency to study perceptual processes in object recognition may find their effects obscured by variance attributable to lexical properties of the object names. An experiment was conducted to determine if reading the object names prior to picture recognition could reduce this variance without interacting with subsequent perceptual processes. Subjects were divided into two groups, one of which read the names of the objects prior to identification and one which did not. Subjects in both groups were required to name object pictures as rapidly as possible. In a first block of trials, subjects identified 16 objects. In a second block, subjects identified 32 objects, half of which were different shaped examples of objects viewed on the first block and half of which were completely new. Significant priming was observed for the different shaped examples in the second block, but not for completely new objects regardless of which group the subject was in. Further, reading the names of objects did not reduce response times or response time variability although it did reduce the number of synonymous name variants subjects used.

##### *Reference*

Cooper, E. E., & Biederman, I. The Effects of Prior Name Familiarization on Object Naming Latencies. Unpublished manuscript.

## II. Neural Net Theory

### A. A Neural Net Implementation of RBC that, More Generally, Offers a Solution to the the Binding Problem.

Upon exposure to a single view of an object, the human can readily recognize that object from any other view that preserves the parts in the original view. Experimental evidence suggests that this fundamental capacity reflects the activation of a viewpoint invariant structural description specifying the object's parts and the relations among them. This paper presents a neural network model of the process whereby a structural description is generated from a line drawing of an object and used for object classification. The model's capacity for structural description derives from its solution to the *dynamic binding problem* of neural networks: Independent units representing an object's parts (in terms of their shape attributes and interrelations) are bound *temporarily* when those attributes occur in conjunction in the systems input. Temporary conjunctions of attributes are represented by synchronized (or phase locked) oscillatory activity among the units representing those attributes. Specifically, the model uses phase locking to: a) parse images into their constituent parts; b) bind together the attributes of a part; and c) determine the relations among the parts and bind them to the parts to which they apply. Because it conjoins independent units temporarily, dynamic binding allows tremendous economy of representation, and permits the representation to reflect the attribute structure of the shapes represented. The model's recognition performance is shown to conform well to empirical findings.

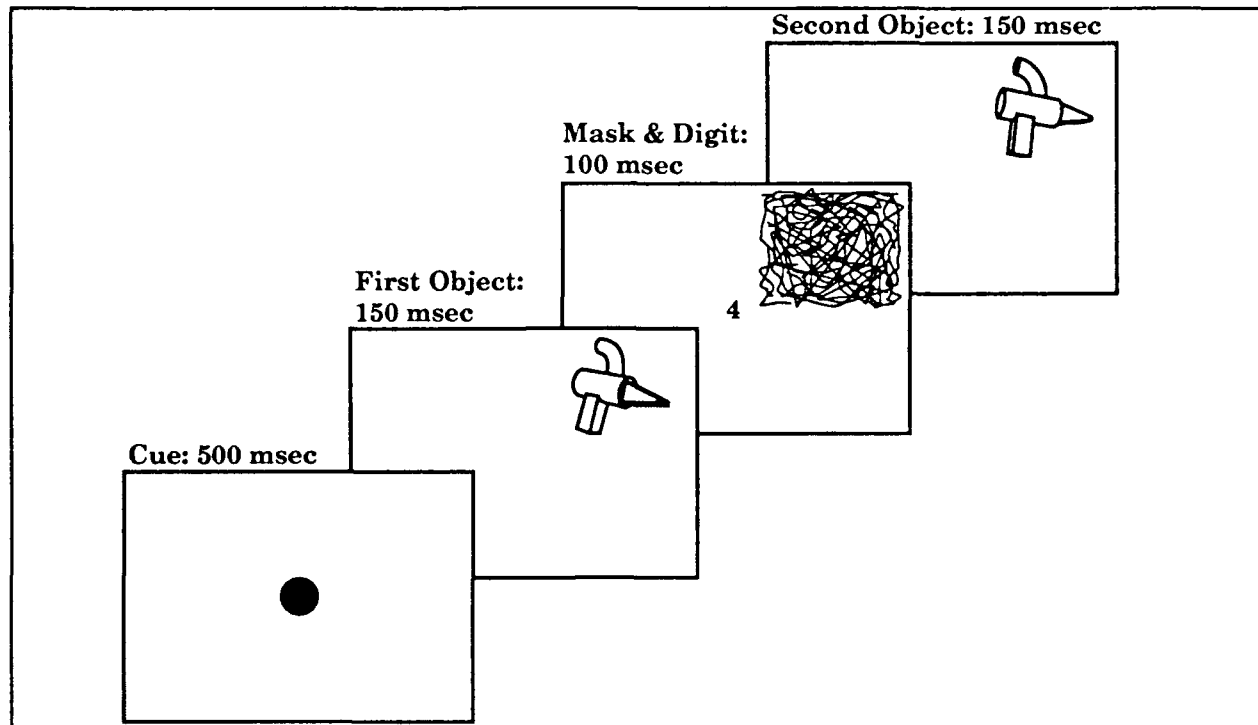
#### Reference

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*. In press.

## III. Cortical Basis of Object Recognition

### A. Object Recognition without a Temporal Lobe

Is the temporal lobe required for high level object recognition? Individuals with one temporal lobe removed (because of seizures) viewed briefly-presented line drawings of objects. The images were presented to the left or right of fixation, so that they would be initially projected to the contralateral hemisphere, and above or below the horizon. The latter feature of the presentation conditions eliminated transfer from V4 to the contralateral temporal lobe though the corpus callosum. Shape information should thus have remained localized to the hemisphere contralateral to the visual field in which the image was shown until the temporal lobe, where rich callosal connections allow transfer to the other temporal lobe in a normal individual. Two kinds of tasks were employed: a) naming (and priming), and b) same-different shape judgments to a sequentially presented pair of pictures, with an intervening mask. In this same-different task, a "same" pair could be identical or rotated up to 60° in depth, as illustrated in Figure 11. "Different" trials used different exemplars with the same name (e.g., two different kinds of chairs). In another same-different, depth-rotation task, nonsense objects composed of simple volumes were used. If processes resident in the temporal lobe are critical for the high level object recognition demanded by these tasks, performance should have been much worse when images were shown in the visual field contralateral to the hemisphere with the missing temporal lobe. Other than a higher error rate for naming images presented to a left hemisphere missing its temporal lobe, differences in performance in recognizing objects presented to a hemisphere with an intact versus absent temporal lobe were minor.



*Figure 11.* Sequence of events on a 60° orientation difference, "same" trial with nonsense objects in the same-different task. Subjects reported central fixation digit after they made their same-different response. This paradigm was also run with familiar objects, where the different trials had same name, different shaped exemplars.

#### *Reference*

Biederman, I., Gerhardstein, P. C., Cooper, E. E., & Nelson, C. A. High level object recognition without a temporal lobe. Abstract submitted for presentation at ARVO, 1992.

### **B. Object Recognition and Laterality: Null Effects**

In two experiments, normal subjects named briefly presented pictures of objects that were shown either to the left or to the right of fixation. The net effects attributable to hemifield were negligible: Naming RTs were 12 msec lower for pictures shown in the left visual field but error rates were slightly lower, by 0.8%, for pictures shown in the right visual field. In both experiments, a second block of trials was run to assess whether hemifield effects would be revealed in object priming. Naming RTs to same name-different shaped exemplar pictures were significantly longer than RTs for identical pictures, thus establishing that a component of the priming was visual, rather than only verbal or conceptual, but hemifield effects on priming were absent. Allowing for the (unlikely) possibility that variables with large differential left-right hemifield effects may be balancing and cancelling each other out, we conclude that there are no differential hemifield effects in either object recognition or object priming.

#### *Reference*

Biederman, I., & Cooper, E. E. (1991). Object recognition and laterality: Null Results. *Neuropsychologia*, 29, 685-694.

### C. Unexceptional Spatial Memory in an Exceptional Memorist

Rajan Mahadevan evidences an exceptional memory for arrays of digits. We tested whether Rajan's spatial memory was likewise exceptional. 8 control Ss and Rajan were instructed to remember the position and orientation of 48 images of common objects shown either to the left or the right of fixation and facing either left or right. Rajan's accuracy for judging whether the position and orientation of these pictures had changed when they were shown in a different sequence was lower than that of control subjects for both judgments. Rajan's exceptional memory capacity apparently does not extend to spatial relations.

#### Reference

Biederman, I., Cooper, E. E., Fox, P. W., & Mahadevan, R. S. (1992). Unexceptional spatial memory in an exceptional memorist. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press.

### D. Lack of Attentional Costs in Detecting Visual Transients.

Both spotlight or zoom-lens metaphors of attention predict that performance should improve at an attended position, and that this advantage should decrease as the area attended increases. These assumptions were tested in a simple detection task (presence or absence of an X) and a simple judgment task (discriminating a dim from a bright X) in different blocks of trials. The target could appear at: a) one of two positions, three degrees to the left or right of fixation, b) one of two positions, six degrees to the left or right of fixation, or c) one of four positions, three or six degrees to the left or right of fixation. Although the experiment was sufficiently sensitive to detect a change in performance due to a modest variation in target luminance, no effect of either eccentricity or number of possible display positions, on either detection or discrimination, was found. The lack of an effect of the number of possible display positions seemed paradoxical given previous research (e.g., Posner, Nissen, & Ogden, 1978) showing a benefit of cueing a position. A third experiment compared performance on the two-location detection task to performance on the same task when subjects were cued as to which of the positions was three times more likely to have a target than the other. Reaction times to targets at the 75% probable position, though shorter than those in the 25% probable condition, were significantly greater than in a 50% probable condition, where subjects received no position cue. The results suggest that detection of targets in the periphery can occur in parallel, without an increase in reaction time as the number and area of possible target locations doubles. Further, the overhead associated with the allocation of attention, within the conditions of these experiments, were greater than any hypothesized benefit from knowledge of target locations.

#### Reference

Fisher, B., Biederman, I., Spivey, M., MacDermott, R., & Bridgman, B. Transient Detection Proceeds in Parallel with a Net Cost from Cueing. Submitted for publication.

### E. Methodology

#### Experimental Software

One methodological goal of the project was to develop a software shell for picture perception experiments. A graduate student, Steve Kohlmeyer, supported by the grant wrote a program, *Picture Perception Lab* (PPL) which does this. PPL capitalizes on the sophisticated drawing software available for the Macintosh computer by allowing the user to create the stimulus images using almost any drawing package. Through a series of user-friendly dialog boxes, PPL

enables even non-programmers to develop, quickly and easily, tachistoscope-like vision experiments. The program allows a wide variety of experimental designs. Image exposure durations and subject reaction times are precisely monitored with millisecond timing functions. Each image is drawn on the screen in a single 60 Hz refresh. Kohlmeyer is currently marketing the software as there appears to be widespread interest in it from the community of those investigators who run picture perception experiments on the Macintosh.

Kohlmeyer, S. (1992). Picture Perception Lab: A program for picture perception experiments on the Macintosh II. *Behavior Research Methods, Instruments, & Computers*, in press.

### Experimental Images

Another methodological goal was to develop a set of high quality pictures using Macintosh drawing packages that could be used in a variety of experiments. To this end, Eric Cooper has created 200 pictures. In this set, there are 32 pairs that have the same name but a different shape, such as a grand piano and an upright piano. These pairs allow assessment of the contribution of response and concept priming apart from the contribution of perceptual processes in picture recognition. Peter Gerhardstein has created 60 3D images using Swivel 3D, so they can be shown at arbitrary orientations in depth. In this set of 60, 48 are from 24 pairs of images that have the same names but a different shape. In addition to the familiar object images, Peter has also created 60 nonsense objects created from unfamiliar arrangements of the geons.

### PUBLICATIONS

Biederman, I. (1988). Real-time human image understanding in pilot performance models. In J. I. Elkind, S. K. Card, J. Hochberg, & B. M. Huey (Eds.) *Human Performance Models for Computer-Aided Engineering*. Pp. 126-143. Washington: National Academy Press. Reissued, Academic Press, 1990.

Hummel, J. E., Biederman, I., Gerhardstein, P. C., & Hilton, H. J. (1988). From image edges to geons: A connectionist approach. In D. Touretzky, G. Hinton, & T. Sejnowski (Eds.) *Proceedings of the 1988 Connectionist Models Summer School*. Pp. 462-471. San Mateo, CA: Morgan Kaufmann.

Biederman, I. (1989). Algorithms for Vision. Review of W. Richards & S. Ullman (Eds.) *Image Understanding, 1985-86*. *Contemporary Psychology*, 34, 7-8.

Biederman, I. (1989). The uncertain case for cross cultural effects in pictorial object recognition. *Behavioral and Brain Sciences*, 12, 74-75.

Biederman, I. (1990). Higher level vision. In D. N. Osherson, S. Kosslyn, & J. Hollerbach (Eds.) *An Invitation to Cognitive Science: Visual Cognition and Action, Volume 2*, Pp. 41-72. Cambridge, MA: MIT.

Biederman, I. (1990). Human Visual Pattern Recognition of Medical Images. *Proceedings of the IV Meeting of the SPIE Society for Medical Imaging*, Pp. 2-7. SPIE-The International Society for Optical Engineering: Bellingham, WA.

Hummel, J. E., & Biederman, I. (1990). Dynamic Binding: A Basis for the Representation of Shape by Neural Networks. *Proceedings of the 12th Annual Meeting of the Cognitive Science Society*, 614-621. Hillsdale, NJ: Erlbaum.

- Biederman, I. (1991). Dumb mechanisms doing smart things. Review of B. Cohen & V. Henn (Eds.) *Representation of Three-Dimensional Space in the Vestibular, Oculomotor, and Visual Systems: A Symposium of the Barany Society*, *Contemporary Psychology*, 35, 979-980.
- Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, 23, 393-419.
- Biederman, I., & Cooper, E. E. (1991). Object recognition and laterality: Null Results. *Neuropsychologia*, 29, 685-694.
- Biederman, I., Hilton, H. J., & Hummel, J. E. (1991). Pattern goodness and pattern recognition. In J. R. Pomerantz & G. R. Lockhead (Eds.) *The Perception of Structure* (Ch. 5, Pp. 73-95). Washington, D.C.: APA.
- Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, 20, 585-593.
- Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 121-133.
- Biederman, I., Cooper, E. E., Fox, P. W., & Mehadevan, R. S. (1992). Unexceptional spatial memory in an exceptional memorist. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, in press.
- Biederman, I., Hummel, J. E., Gerhardstein, P. C., & Cooper, E. E. (1992). From Image Edges to Geons to Viewpoint Invariant Object Models: A Neural Net Implementation. *Proceedings of the International Society for Optical Engineering Conference on Intelligent Information Systems*, In press.
- Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*. In press.
- Kohlmeyer, S. (1992). Picture Perception Lab: A Program for Picture Perception Experiments on the Macintosh II. *Behavior Research Methods, Instruments, & Computers*, in press.
- Biederman, I. (1992). Human Image Understanding. In P. Johansen (Ed.) *Image Analyses*. Singapore: World. In press.
- Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: A review and additional evidence. *Canadian Journal of Psychology*, In press.
- Biederman, I., Hummel, J. E., Cooper, E. E., & Gerhardstein, P. C. (1992). Shape recognition in mind, brain, and machine. In P. Rudomen & F. Cervantes (Eds.) *Natural and Artificial Intelligence: A Meeting Between Neuroscience and Artificial Intelligence*, In press.
- Cooper, E. E., & Biederman, I. (1992). Priming objects with single volumes: Searching for the representational locus of perceptual priming. Submitted to *ACTA Psychologica*.
- Biederman, I., Gerhardstein, P. C., Cooper, E. E., & Nelson, C. A. (1992). High level object recognition without an inferior temporal lobe. Submitted to the *Journal of Cognitive Neuroscience*.

Biederman, I., & Gerhardstein, P. C. (1992). Recognizing depth-rotated objects: Evidence for 3D viewpoint invariance. Submitted to the *Journal of Experimental Psychology: Human Perception and Performance*.

Fisher, B., Biederman, I., Spivey, M., MacDermott, R., & Bridgman, B. Transient Detection Proceeds in Parallel with a Net Cost from Cueing. Submitted for publication.

## PRESENTATIONS AT SCIENTIFIC MEETINGS

Biederman, I. (1988). Invariant Visual Primitives for Object Recognition. Invited address to the AAAS Symposium on High-Level Vision; Interdisciplinary Approaches to Object Recognition. Boston, Massachusetts: February.

Biederman, I. (1988). Geons: Visual Primitives for Human Image Understanding. Invited paper presented at the Meeting for Visual Form and Motion Perception: Psychophysics, Computation, and Neuro Networks. Boston, Massachusetts: March.

Biederman, I. (1988). Real-time human image understanding. Paper presented at the Spring Symposium on Physical and Biological Approaches to Computational Vision, American Association for Artificial Intelligence. Stanford, CA.: March.

Biederman, I. (1988). Human image understanding. Invited address to the Meetings of the Midwestern Psychological Association. Chicago, IL.: April.

Biederman, I. (1988). Visual object recognition. Invited presentation at the James S. McDonnell Foundation Summer Institute in Cognitive Neuroscience. Harvard University: June.

Biederman, I., Blicke, T. W., Ju, G., Hilton, H. J., & Hummel, J. E. (1988). Empirical analyses and connectionist modeling of real-time human image understanding. Invited presentation at the Tenth Annual Conference of the Cognitive Science Society, Montreal, Canada, August.

Hummel, J. E., Biederman, I., Gerhardstein, P., & Hilton, H. J. (1988). A connectionist model of Recognition-by-Components. Paper presented at the Neural Architectures for Computer Vision Workshop at the Meetings of the Seventh National Conference on Artificial Intelligence. St. Paul, MN. (August).

Biederman, I. (1988). Real-time human image understanding. Invited address to the XXIV International Congress of Psychology. Sydney, Australia: August.

Biederman, I. (1989). Pattern goodness and pattern recognition. Invited presentation at a Conference on the Perception of Structure. Yale University, May.

Biederman, I. (1989). Viewpoint invariant primitives as a basis for human object recognition. Invited presentation at a Conference on Vision and Three-Dimensional Representation. University of Minnesota, Minneapolis, MN: May.

Biederman, I., Cooper, E. E., and Hilton, H. J. (1989). Priming with contour-deleted images. Invited presentation at the Symposium on *Issues in Object Recognition: Perception, Attention, and Memory* to be presented at the First Meeting of the American Psychological Society, Alexandria, VA. June 12, 1989.



- Biederman, I. (1989). Volumetric models for object recognition: Some implications for theories of extrastriate processing. Invited paper at the Canadian Institute for Advanced Research Workshop on Extrastriate Visual Computation. Montreal, June 13-14.
- Biederman, I., Cooper, E. E., & Gerhardstein, P. C. (1989). Priming fragmented images. Paper presented at the 30th Annual Meeting of the Psychonomic Society, Atlanta, GA, November 17-19.
- Biederman, I. (1989). What is there left to learn from the existence proof for 3D pattern recognition? Invited presentation at the IEEE Workshop on the Interpretation of 3D Scenes. Austin, Tx., November.
- Biederman, I. (1990). Synchrony of firing in neural net architectures: A common source for perceptual organization and attentional limitations? Invited paper at the National Research Council Committee of Vision Conference on Visual search: Segmentation, Identification, and Attention. Newport Beach, CA, January 19-21.
- Biederman, I. (1990). Human Visual Pattern Recognition. Invited address at the Meetings of the SPIE Society for Medical Imaging. Newport Beach, CA. February.
- Biederman, I., & Cooper, E. E. (1990). Evidence for intermediate, translation invariant representations in visual object recognition. Poster presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.
- Hummel, J. E., & Biederman, I. (1990). A neural net implementation of Recognition-by-Components. Poster presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.
- Fisher, B., Bridgeman, B., & Biederman, I. (1990). Task differences in visual search: Does attention aid detection? Paper presentation at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.
- Biederman, I., & Hummel, J. E. (1990). Grounding mental symbols in object images. Invited paper presented at the Symbol Grounding Workshop at the 16th Annual Meeting of the Society for Philosophy & Psychology, University of Maryland, College Park, Md. June.
- Biederman, I., Hummel, J. E., & Cooper, E. E. (1990) Human Object Recognition. Invited address presented at a *Conference on Visual Information Assimilation in Man and Machine*, Ann Arbor, Michigan, June.
- Hummel, J. E., & Biederman, I. (1990). Dynamic Binding: A Basis for the Representation of Shape by Neural Networks. Paper presented at the 12th Annual Meeting of the Cognitive Science Society, Cambridge, MA. July.
- Biederman, I., & Cooper, E. E. (1990). Intermediate, invariant representations mediate visual object recognition. Invited presentation at a Workshop on Object and Scene Perception. University of Leuven, Leuven, Belgium. September.
- Hummel, J. E., & Biederman, I. (1990). Binding invariant shape descriptors: A neural net architecture for structural description and object recognition. Invited presentation at a Workshop on Object and Scene Perception. University of Leuven, Leuven, Belgium. September.

- Biederman, I. (1990) Visual Image Understanding. *The Fourth Annual Fern Forman Fisher Lecture*, University of Kansas, November.
- Hummel, J. E., & Biederman, I. (1990). Binding invariant Shape Descriptors for Object Recognition: A Neural Net Implementation. Paper presented at the 31th Annual Meeting of the Psychonomic Society. New Orleans, LA, November 17-19.
- Biederman, I. (1991). How an account of Shape Recognition can be Achieved by a Neural Network that Solves the Binding Problem through Phase Locking. Invited paper presented at the Workshop on Rhythmic Oscillations in Cortex: Their Form and Function, Tucson, April.
- Hummel, J. E., & Biederman, I. (1991). Binding by phase locked neural activity: Implications for a theory of visual attention. Paper presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.
- Cooper, E. E., & Biederman, I. (1991). Evidence for size invariant representations in visual object recognition. Poster presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.
- Gerhardstein, P. C., & Biederman, I. (1991). 3D Orientation invariance in visual object recognition. Paper presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, Fl. May.
- Biederman, I. (1991). The neuroscience of object recognition. Invited featured speaker at The First Annual Meeting of the Canadian Society for Brain, Behavior, and Cognitive Science, Calgary, June.
- Biederman, I. (1991). Shape recognition in eye and brain. Invited presentation at the Stockholm Workshop on Computational Vision, Rosenen, Sweden, August.
- Biederman, I. (1991). Human Image Understanding. Invited address presented at the 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark. August.
- Biederman, I., Cooper, E. E., & Gerhardstein, P. C. (1991). Picture naming reveals the major invariances expected of a shape recognition system. Poster presented at the Meetings of the Psychonomic Society, San Francisco, CA. November.
- Mansfield, J. S., Biederman, I., Legge, G. E., & Knill, D. C. (1991). Greater statistical efficiency for viewpoint-invariance differences in the categorization of curves. Paper presented at the Meetings of the Optical Society, San Jose: CA. November.
- Biederman, I. (1992). The neural basis of shape recognition. Invited paper presented at the Meetings of the AAAS, Chicago, February, 1992.
- Biederman, I. (1992). Shape recognition in Mind and Brain. Invited paper presented to the Helmholtz Society, University of California at Irvine, February, 1992.
- Biederman, I. (1992). Reverse Engineering the Psychology of Shape Recognition. Invited paper presented at the Office of Naval Research Workshop on Intermediate and Higher Level Vision. Laguna Beach, CA. March.
- Biederman, I., & Hummel, J. E. (1992). From Image Edges to Geons to Viewpoint Invariant Object Models: A Neural Net Implementation. Invited paper to be presented at the International

Society for Optical Engineering Conference on Intelligent Information Systems, Orlando, FL.  
April.

Biederman, I., Gerhardstein, P. C., Cooper, E. E., & Nelson, C. A. (1992). High level object recognition without a temporal lobe. Poster to be presented at the Annual Meeting of The Association for Research in Vision and Ophthalmology, Sarasota, FL. May.

Biederman, I. (1992). Human Image Understanding. Invited paper to be presented at the Meetings of the International Congress of Psychology, Brussels, Belgium, August, 1992.

## INVITED COLLOQUIA

MIT (Psychology Department and the Center for Biological Information Processing)  
US Army Research Institute, Ft. Benning, Ga  
Stanford University  
McGill University (Departments of Electrical Engineering; Cognitive Science Program)  
University of Auckland, New Zealand  
Florida Atlantic University  
Hebrew University of Jerusalem  
University of Haifa  
Ben Gurion University of the Negev  
Weitzmann Institute (Department of Computer Science)  
University of Tel Aviv  
Centre National de la Recherche Scientifique (Paris)  
Ecole National Scientifique Technical (Paris)  
University of Nijmegen, The Netherlands  
University of Leuven/Louvain, Belgium  
North Dakota State University  
Princeton University  
National Institutes of Health (Neuropsychology Laboratories)  
George Washington University (Computer Science Department)  
University of Arizona (Psychology Department and Cognitive Science Program)  
University of Southern California (Department of Psychology; Department of Computer Science);  
University of Paris  
University of Kansas  
Rice University  
University of California, Berkeley  
Stanford University  
University of California, Santa Cruz  
Navel Ocean Systems Center, Kailua Bay, Hawaii  
University of Hawaii at Manoa  
MIT Cognitive Neuroscience Program  
Indiana University  
University of Pennsylvania, Computer Science Department  
Pennsylvania State University.  
University of Toronto  
California Institute of Technology  
UCLA

## HONORS

Offered (and accepted) the William M. Keck Endowed Chair of Psychology at the University of Southern California

Invited to be a Fellow at the Center for Advanced Study in the Behavioral Sciences, Palo Alto, California.

Elected to the Helmholtz Club.

Invited to present the Fourth Fern Forman Fisher Lecture, University of Kansas

Nominated for Editorship of *Journal of Experimental Psychology: Human Perception and Performance* (declined).

Nominated for Editorship of *Journal of Experimental Psychology: Learning, Memory, & Cognition* (declined).

Member, *National Research Council Committee on Vision*

Member, National Science Foundation, Science and Technology Centers Panel, 1989

Co-organizer, Workshop on Object and Scene Perception, Leuven, Belgium, 1990.